SYSTEM AND METHOD FOR PROVIDING AN INTERACTIVE, VISUAL COMPLEMENT TO AN AUDIO PROGRAM

This application claims the benefit of U.S. Provisional Patent Application 60/315,046, filed on August 28, 2001.

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention is generally related to audio services, and, more specifically, provides an interactive, visual complement to one or more audio programs.

2. Discussion of the Background

Presently, there exist systems that broadcast music via satellite and cable to consumers' televisions or set-top boxes or other broadcast receiving devices. Within such a system, a consumer has typically a selection of 45 music channels to choose from. The channels comprise a variety of music genres and formats. Conventionally, for each of the available music channels, the system broadcasts audio only or, at most, a few lines of text in addition to the audio. This additional text is displayed on the consumer's TV screen. On any given channel, the text typically includes information about the music that is currently playing on that channel, such as the name of the artist, the title of the song, and the title of an album that contains the song.

Because only a few lines of text, at most, are transmitted with the audio, a consumer who tunes his or her TV or set-top box to one of the music channels sees an almost entirely blank TV screen. Thus, in conventional broadcast music systems, the TV screen is underutilized and the consumer's overall enjoyment of the system is

limited.

What is desired, therefore, is a system to overcome this and other disadvantages of conventional music systems.

SUMMARY OF THE INVENTION

5 The present invention overcomes the above described disadvantage by providing a system and method for providing a visual complement to one or more audio programs. In one aspect, the system includes an audio subsystem for selecting a sound recording based on a playlist, generating an audio signal corresponding to the sound recording, and transmitting triggers to a video subsystem whenever a sound recording is selected.

10 Upon receiving a trigger from the audio subsystem, the video subsystem generates a video image specification based, at least in part, on the selected sound recording. The audio signal and video image specification are transmitted to an audio/video signal transmission system. The transmission system receives the video image specification and generates a video image that conforms to the video image specification. The

15 transmission system then transmits the video image and the audio signal to consumers' audio/video receivers so that the audio signal and video image may be perceived by the consumers. In this way, the system provides a visual complement to an audio service.

In one embodiment, the audio/video signal transmission system is a broadcast transmission system that broadcasts the video image and the audio signal to the

20 consumers' audio/video receivers.

Advantageously, the invention may also provide an interactive, visual complement to the audio program. In this embodiment, the transmission system adds one or more selectable, interactive buttons to the video image depending on information

received from the video subsystem.

In another aspect, the system also includes a video image generator coupled to the video subsystem. In this aspect, the video image specification generated by the video subsystem in response to the trigger received from the audio subsystem is provided to the video image generator. The video image generator then generates a video image based on the provided video image specification and transmits the video image to a first transmission subsystem. At the same time this is occurring, audio subsystem transmits the audio signal corresponding to the selected sound recording to the first transmission subsystem. The first transmission subsystem then transmits the audio signal together with the video image to a second transmission system, which then transmits the audio signal and video image to the consumers' receivers so that when a consumer tunes his receiver to the particular channel the consumer will be able to hear the sound recording and view the video image.

Advantageously, the video image is updated at various times so that the video image seen by the consumer changes over time as well as changing whenever a new sound recording is selected and played by the audio subsystem.

In one particular aspect, the video subsystem generates an HTML document that contains the video image specification and provides the HTML document to the video image generator. The video image generator uses the HTML document to generate an MPEG video presentation.

In another aspect, the video images are pre-generated. The pre-generated video images may be stored at the audio/video system or at the transmission system. Advantageously, a data structure is used to associate a set of one or more of the pre-

generated video images with one or more sound recordings from a playlist.

Further features and advantages of the present invention, as well as the structure and operation of various embodiments of the present invention, are described in detail below with reference to the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings, which are incorporated herein and form part of the specification, illustrate various embodiments of the present invention and, together with the description, further serve to explain the principles of the invention and to enable a person skilled in the pertinent art to make and use the invention. In the drawings, like reference numbers indicate identical or functionally similar elements. Additionally, the left-most digit(s) of a reference number identifies the drawing in which the reference number first appears.

FIG. 1 is a block diagram of one embodiment of an audio/video system for providing audio/video programming to consumers.

FIG. 2 illustrates various locations on a TV screen where visual media assets may be displayed.

FIGS. 3A-3C are flow charts illustrating processes, according to one embodiment, performed by the audio subsystem, the video subsystem, and the audio/video signal transmission system, respectively, for providing an interactive, visual complement to an audio program for a particular channel.

FIG. 4 illustrates pre-defined configuration data that is associated with a particular channel and that is used by the video subsystem to create data packets for the particular channel.

FIGS. 5A and 5B is a flow chart illustrating a process, according to one embodiment, for creating a data packet for a particular channel.

FIG. 6 illustrates an exemplary data packet.

FIG. 7 is a block diagram of a system according to another embodiment of the invention.

FIG. 8 is a flow chart illustrating a process, according to another embodiment, that is performed by the video subsystem.

FIG. 9 is a flow chart illustrating a process, according to one embodiment, that is performed by the video image generator.

FIG. 10 is a block diagram of a system according to another embodiment of the invention.

FIG. 11 is a flow chart illustrating a process, according to one embodiment, that is performed by the video subsystem.

FIG. 12 illustrates an exemplary data structure that associates sound recording identifiers from a playlist with a set of one or more video image identifiers.

FIG. 13 is a flow chart illustrating a process, according to one embodiment, that is performed by the audio/video signal transmission system 170 when the video images are pre-generated.

FIG. 14A is a flow chart illustrating a process, according to one embodiment, that is performed by the video subsystem when the video images are pre-generated.

FIG. 14B is a flow chart illustrating a process, according to another embodiment, that is performed by the audio/video signal transmission system when the video images are pre-generated.

FIG. 15A is a flow chart illustrating a process, according to another embodiment, that is performed by the video subsystem when the video images are pre-generated.

FIG. 15B is a flow chart illustrating a process, according to another embodiment, that is performed by the audio/video signal transmission system when the video images are pre-generated.

FIG. 16 is a flow chart illustrating a process, according to another embodiment, that is performed by the video subsystem when the video images are pre-generated.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

FIG. 1 is a block diagram of one embodiment of a system 100 for providing audio/video programming. System 100 includes an audio/video system 101 comprising an audio subsystem 102 that provides audio content for transmission to listeners over one or more channels and a video subsystem 104 for providing video content that is transmitted together with the audio content and that complements the audio content. System 100 further includes a transaction processing subsystem 106 for processing transactions, such as electronic commerce ("e-commerce") transactions.

Audio/video system 101 may comprise a data processing system, a persistent storage device, and volatile memory. Stored in the storage device and/or the volatile memory are computer instructions (i.e., software) that enable audio/video system 101 to perform the functions and processes described herein. Audio subsystem 102 and video subsystem 104 may be implemented in software or a combination of software and hardware.

Audio subsystem 102 has access to a sound recording library 105 that includes a

large number of sound recordings (e.g., tracks from albums of many different genres). The sound recordings may be stored on compact discs, hard disks, or other media for storing data.

Audio subsystem 102 preferably includes a playlist 110 for each of the one or more channels supported by system 100. A playlist 110 for a particular channel specifies sound recordings that have been programmed for transmission to the listeners of system 100 over that channel during a given period of time. A new playlist 110 is typically generated for each channel on some periodic basis (e.g., daily, weekly, etc.).

Audio subsystem 102 typically retrieves, encodes, and streams the sound recordings to consumers in the order in which the sound recordings are listed in the playlists 110. Preferably, the sound recordings are encoded by audio subsystem 102 according to the Dolby AC-3 coding technique.

Audio subsystem 102 may stream the encoded sound recordings to a transmission subsystem 190, which may transmit the encoded sound recordings to an audio/video signal transmission system 170. Transmission system 170 may be a broadcast transmission system, such as a cable head-end or a direct broadcast satellite system. Transmission system 170 comprises a transmitter (not shown) for transmitting signals and a computer (not shown) programmed to perform processes described herein.

Transmission system 170 transmits the encoded sound recordings to audio/video receivers 180, which are coupled to an audio/video device 182 that reproduces the sound recordings for the subscribers. Receivers 180 may be conventional digital cable or satellite set-top boxes. Audio/video device 182 may comprise a TV screen or monitor and speakers.

Video subsystem 104, in one embodiment, is responsible for, among other things, generating, in real time, data packets for each of the one or more channels. A data packet for a particular channel comprises a video image specification that specifies a visual complement of the audio service for the particular channel. Thus, the video image specification defines how the listeners' TV screens will look when the listener tunes to the particular channel.

More specifically, the video image specification specifies one or more visual media asset identifiers, each of which identify one or more visual media assets. The video image specification may also specify the screen position where each identified asset is to be displayed. Examples of video media assets include: graphic image files (e.g., GIF files, JPEG files, bitmap files, etc.), video files (e.g., MPEG files, AVI files), text messages, etc. It is these assets that are used to create the visual complement to the audio service.

The video image specification for a particular channel is based, at least in part, on the sound recording that the particular channel is currently playing. Therefore, for example, if a U2 song from the Joshua Tree album is currently being played on channel 51, then, at some particular point in time while the song is playing, the video image specification for channel 51 might specify that an image of the Joshua Tree album art is to be displayed at a first location 202 (see FIG. 2) on a TV screen (or monitor) 282.

Additionally, the video image specification may also specify that the name of the song, artist, and album is to be displayed at a second location 204 on the TV screen 282, and an advertising banner is to be displayed at a third location 206 on the TV screen 282.

In one embodiment, the video image specification may also specify that certain music trivia and/or news is to be displayed at a fourth location 208 on the TV screen 282. It should be understood that album art, advertising banners, text messages, and other visual media assets may be positioned anywhere on the TV screen 282 and that the invention is not limited to the particular arrangement of visual media assets shown in FIG. 2.

The video image specification may also be time driven. That is, at least some of the assets (e.g., advertising banners, music trivia, and news) specified by the video image specification are determined as function of time, regardless of which sound recording is currently playing.

Preferably, each video image specification for a particular channel includes an asset identifier that identifies a text message that contains information pertaining to the sound recording that is currently being played over the particular channel. This information may include the name of the artist who created the sound recording, the title of the sound recording, and the name of an album on which the sound recording can be found. Alternatively, instead of or in addition to each video image specification for the particular channel including the asset identifier that identifies the text message, the text message itself may be included in the data packet.

In addition to including a video image specification, the data packet may further include purchase information for enabling a listener of system 100 to purchase the album or the sound recording. The purchase information may include an indicator that the sound recording or album is saleable, a price, and a unique code that identifies the album.

FIG. 6 illustrates an exemplary data packet 600. As shown in FIG. 6, data packet 600 includes a video image specification 602. Optionally, data packet 600 may also include sound recording information 604, and purchase information 606. Video image specification 602 comprises a list of visual media asset identifiers and associates a screen position with each asset identifier. The data packets may be extensible mark-up language (XML) files or hyper-text mark-up language (HTML) files.

In the embodiment shown in FIG. 1, after generating a data packet for a particular channel, video subsystem 104 transmits the data packet so that it will be received by transmission system 170. Video subsystem 104 may use transmission subsystem 190 to transmit the data packet to transmission system 170 or may use a public network (e.g., the Internet) or private network to transmit the data packet to transmission system 170.

Transmission system 170 may have access to a data storage unit 185. Preferably, storage unit 185 has a very short access time. Storage unit 185 stores the visual media assets specified in the data packet (storage unit 185 is updated periodically by an administrator to ensure that storage unit 185 contains the necessary visual media assets). Therefore, borrowing from the above example, storage unit 185 stores the image of the Joshua Tree album art that is displayed when the song from U2's Joshua Tree album is playing.

In embodiments where transmission system 170 does not have access to storage unit 185, a storage unit 186 that is coupled to video subsystem 104 stores the visual media assets specified in the video image specification, and video subsystem 104 retrieves the assets from storage 186 and transmits them to transmission system 170.

After receiving the data packet for the particular channel, transmission system 170 parses the data packet and determines the video image specification and purchase information that are specified therein. Transmission system 170 then creates a video image corresponding to the video image specification and transmits the video image over the particular channel to subscribers' audio/video receivers 180. The video image is then displayed by audio/video device 182.

The video image conforms to the video image specification contained in the data packet so that when the video image is displayed on the subscribers' audio/video device 182, the visual media assets defined in the video image specification are displayed in the locations as specified in the video image specification.

The video image may be encoded according to a Moving Pictures Experts Group (MPEG) standard, the National Television Standards Committee (NTSC) video signal standard, or other video signal standard. In one specific embodiment, the video image is encoded according to an MPEG standard and comprises an MPEG I-frame followed by null P-frames.

FIGS. 3A-3C are flow charts illustrating processes 300, 330, and 360, according to one embodiment, performed by audio subsystem 102, video subsystem 104, and transmission system 170 respectively, for providing an interactive, visual complement to the audio service for a particular channel. The same process is performed for the other channels.

Process 300 (see FIG. 3A) begins in step 302, where audio subsystem 102 selects a sound recording from library 105 based on a playlist for the particular channel. After selecting the sound recording, audio subsystem 102 retrieves it from library 105,

encodes it, and transmits it to transmission subsystem 190 (step 304), which then transmits it to a system, such as, for example, a transmission system 170, that transmit audio/video signals to the subscribers' receivers 180.

At or about the same time as step 304 is performed, audio subsystem 102 transmits to video subsystem 104 a trigger message specifying a sound recording identifier that identifies the sound recording selected in step 302, sound recording information pertaining to the sound recording, and a channel identifier (step 306). The sound recording identifier uniquely identifies the sound recording selected in step 302 and the channel identifier uniquely identifies the particular channel. After audio subsystem 102 finishes transmitting the sound recording selected in step 302, control passes back to step 302, where audio subsystem 102 selects another sound recording from library 105 based on the playlist for the particular channel after it has finished streaming the previously selected sound recording for that channel.

Process 330 (see FIG. 3B) begins in step 332, where video subsystem 104 waits for a trigger message from audio subsystem 102 or for a timer to expire. If video subsystem 104 receives a trigger message from audio subsystem 102, control passes to step 334, and if a timer expires, control passes to step 338.

In step 334, video subsystem 104 parses the trigger message to determine the sound recording identifier, sound recording information, and channel identifier specified therein. Next (step 336), video subsystem 104 uses this information, together with pre-defined configuration data that is associated with the channel identified by the channel identifier, to create a data packet for the identified channel. The predefined configuration data is stored in video subsystem 104. An illustration of pre-defined

configuration data is shown in FIG. 4, and will be discussed in more detail further

below.

In step 338, video subsystem 104 determines a channel and an asset identifier

queue that is associated with the expired timer (see element 420 of FIG. 4 for an

illustration of an exemplary queue). Next (step 340), video subsystem 104 may create a

data packet for the identified channel based, at least in part, on the contents of the asset

identifier queue associated with the expired timer. An illustration of a process 500 for

creating a data packet is shown in FIG. 5, and will be discussed in more detail further

below.

After creating the data packet in either step 336 or 340, video subsystem 104

transmits the data packet to audio/video transmission system 170 (step 342). After step

342, control passes back to step 344. In step 344, video subsystem retrieves from a

storage unit 186 the visual media assets specified in the data packet and transmits the

assets to transmission system 170 if storage unit 185 does not contain the assets.

Process 360 (see FIG. 3C) begins in step 362. In step 362, audio/video signal

transmission system 170 receives from transmission subsystem 190 the audio stream

transmitted by audio subsystem 102. Next (step 364), transmission system 170

transmits the audio stream to receivers 180.

While transmission system 170 is receiving and transmitting the audio stream,

transmission system 170 receives from video subsystem 104 a data packet for the

particular channel (step 366). After receiving the data packet for the particular channel,

transmission system 170 parses the data packet and determines the video image

specification and purchase information (if any) specified therein (step 368). That is,

transmission system 170 determines the set of asset identifiers specified by the video image specification and the screen position associated with each asset identifier, which may also be specified by the video image specification.

Next (step 370), transmission system 170 retrieves from storage unit 185 the assets identified by the asset identifiers determined in step 368, but if storage unit 185 does not have the assets, then transmission system 170 receives them from video subsystem 104, as described above.

Next (step 372), transmission system 170 determines whether the purchase information indicates that a "Buy" button 250 and/or "Buy-Previous" button 251 should be included in of the video image transmitted to receivers 180. Buy button 250 and Buy-Previous button 251 are interactive, selectable buttons that a user of system 100 may select if the user desires to make a purchase.

If it is determined that Buy button 250 and/or Buy-Previous button 251 should be included in the video image transmitted to receivers 180, then control passes to step 374, otherwise control passes to step 376.

In step 374, transmission system 170 uses the assets retrieved in step 370 and screen position information determined in step 368 to create a video image that conforms to the video image specification contained in the data packet. In step 376, transmission system 170 performs the same step as in step 374, but also adds Buy button 250 and/or Buy-Previous button 251 to the video image. After step 374 and step 376, control passes to step 378. In step 378, the video image created in step 374 or 376 is transmitted to receivers 180. After step 378, control passes back to step 366.

Alternatively, transmission system 170 does not perform step 376. Rather, if it

is determined that Buy button 250 and/or Buy-Previous button 251 should be included in the video image created in step 372, then transmission system 170 sends one or more commands to receivers 180 that direct the receivers 180 to overlay Buy button 250 and or Buy-Previous button 251 onto the vide image transmitted in step 378, provided that receivers 180 are capable of overlying selectable buttons.

A listener who desires to purchase a saleable item may select the Buy 250 or Buy-Previous 251 button to initiate a conventional e-commerce transaction with transaction processing system 106. The listener may select the Buy or Buy-Previous button by, for example, selecting a pre-defined button on a remote control (not shown) that communicates with a receiver 180.

In response to the listener selecting a button 250 or 251, a user interface screen is presented on audio/video device 182. The screen provides information regarding the product (i.e., the album or song currently playing), such as purchase price. If the listener decides to purchase the product, the listener may, for example, select another pre-defined button on the remote control. This will cause a message to be sent from the listener's receiver 180 to transaction processing system 106. The message indicates that the listener desires to purchase the product and may contain an identifier that identifies the product and an identifier that identifies the listener or a registered user account. The receiver may directly send the message to the system 106 through a network, such as the Internet, or may send the message to transmission system 170, which then relays the message to system 106. Upon receiving the message, transaction processing system 106 process the purchase transaction and/or communicates with a vendor who provides the product.

Referring now to FIG. 4, FIG. 4 illustrates pre-defined configuration data 400 that is associated with a particular channel and that is used by video subsystem 104 to create data packets for the particular channel. As shown in FIG. 4, the pre-defined configuration data 400 associates visual media asset identifiers with sound recording identifiers. Each asset identifier uniquely identifies a visual media asset. Thus, configuration data 400 associates visual media assets with a sound recordings.

Preferably, the visual media assets associated with a sound recording are to be displayed during the entire time the sound recording is being played. For example, as shown in FIG. 4, sound recording identifier 402 is associated with asset identifiers 404 and 406. Thus, when system 100 plays the sound recording identified by sound recording identifier 402, the assets identified by asset identifiers 404 and 406 should be displayed to the listeners. Preferably, the configuration data associates a position with each visual media asset. For example, assets 404 and 406 are associated with positions 5 and 3 respectively.

The configuration data may also specify one or more asset queues. An asset queue is an ordered list of asset identifier sets. An asset identifier set contains one or more asset identifiers and a screen position for each asset identifier. Preferably, a time duration is associated with each asset identifier set in a queue. For example, the exemplary configuration data 400 illustrated in FIG 4, specifies two asset queues: queue 420 and 430. Queue 420, for example, contains asset sets 421-423, and assets 421-423 are associated with a time duration of 30 seconds, 60 seconds, and 45 seconds, respectively. As an example, asset identifier set 421 contains asset identifiers 491 and 492, where asset identifier 491 is associated with screen position 1 and asset identifier

492 is associated with screen position 2.

In addition to associating a sound recording identifier with certain asset identifiers, the configuration data may also associate a sound recording identifier with one or more of the asset identifier queues. For example, as shown in FIG. 4, sound recording 402 is associated with asset identifier queue 420 and 430. Because asset sets 421-423 are listed in queue 420 and because queue 420 is associated with sound recording 402, assets identified by asset identifier sets 421-423 are displayed sequentially for the specified duration of times while sound recording 402 is being played. That is, while sound recording 402 is being played, the assets identified by asset identifier set 421 are displayed for its specified duration (i.e., 30 seconds), followed by the assets identified by asset identifier set 422 for its specified duration (i.e., 60 seconds), and then followed by the assets identified by asset identifier set 423 for its specified duration (i.e., 45 seconds).

Referring now to FIGS. 5A and 5B, FIGS. 5A and 5B is a flow chart illustrating a process 500, according to one embodiment, for creating a data packet for a particular channel. Process 500 begins in step 501 wherein video subsystem 104 initializes a data packet so that it does not contain any data. Next (step 502), video subsystem 104 determines whether a trigger message from audio subsystem has been received. If a trigger message is received, control passes to step 504, otherwise control passes to step 503. In step 503, video subsystem 104 determines whether an asset queue timer has expired. If an asset queue timer expires, control passes to step 540, otherwise control passes back to step 502.

In step 504, video subsystem 104 parses the trigger message to determine the

sound recording identifier, sound recording information, and channel identifier specified therein. Next (step 506), video subsystem 104, uses the pre-defined configuration data to determine a set of assets identifiers that are associated with the sound recording identifier last determined in step 504. Video subsystem 104 then

5   determines the screen position that is associated with each asset identifier in the set (step 508). The asset identifiers determined in step 506 and their associated screen positions determined in step 509 are included in the data packet (step 510).

Next (step 512), video subsystem 104 uses the pre-defined configuration data to determine whether there are any asset identifier queues associated with the sound

10   recording identifier determined in step 504. If there are, control passes to step 514, otherwise control passes to step 528.

In step 514, video subsystem 104 selects one of the queues that the configuration data indicates is associated with the sound recording identifier. Next (step 516), video subsystem determines the asset identifier set in the selected queue that is at the "head"

15   of the selected queue. In one embodiment, video subsystem 104 maintains a head pointer for each queue specified by the configuration data. The head pointer for a queue points to the asset identifier set in the queue that is at the head of the queue. Thus, video subsystem 104 may use the head pointer to determine the asset identifier set in the selected queue that is at the head of the selected queue. After step 516, control passes

20   to step 518.

In step 518, video subsystem 104 includes in the data packet each asset identifier listed in the asset identifier set determined in step 516 together with each asset identifier's associated screen position. Next (step 520), video subsystem 104

determines the duration associated with the asset identifier set determined in step 516. Next (step 522), video subsystem 104 activates the timer associated with the selected queue so that the timer will expire after X amount of time has expired, where X is equal to the duration determined in step 518. After step 522, control passes to step 524.

In step 524, video subsystem 104 determines whether there are additional asset identifier queues associated with the sound recording identifier. If there are, control passes to step 526, otherwise control passes to step 528. In step 526, video subsystem 104 selects a queue that is associated with the sound recording and that has not already been selected since the trigger message was received. After step 526, control passes back to step 516.

In step 528, video subsystem 104 includes in the data packet the sound recording information and purchase information included in the trigger message received in step 502. This information concerns the sound recording identified by the sound recording identifier determined in step 504. In one embodiment, the trigger message does not include this information, rather, this information is included in the pre-defined configuration data. More specifically, the pre-defined configuration data associates sound recording information and purchase information with each sound recording identifier included in the configuration data, as shown in FIG. 4. After step 528, control passes to step 530, where the data packet is transmitted to transmission system 170. After step 530, control passes back to step 502.

In step 540, video subsystem 104 determines the queue that is associated with the timer that expired. Next (step 542), video subsystem 104 increments the head pointer associated with the queue determined in step 540 to point to the next asset

identifier set in the queue if the queue determined in step 540 is associated with the

sound recording identifier determined in step 504. However, if the head pointer is

pointing to the last asset identifier set in the queue, video subsystem resets the pointer to

point to the asset identifier set that is at the top of the queue. In this way, the queues are

5   circular queues. After step 542, control passes to step 506.

FIG. 7 is a block diagram of a system 700 for providing audio/video

programming according to another embodiment of the present invention. System 700 is

identical to system 100 with the exception that system 700 further includes a video

image generator 702 that is coupled to video subsystem 104. Video image generator

10   702 has access to storage 186, which stores the visual media assets necessary to create

the visual complement to the audio service.

Additionally, instead of transmission system 170 receiving data packet 131

generated by video subsystem 104, as described above with respect to FIG. 1, video

image generator 702 receives a data packet 732 generated by video subsystem 104.

15   Data packet 732 comprises a video image specification. Further, video subsystem 104

may also generate a data packet 731 and transmits data packet 731 to transmission

subsystem 190. Data packet 731 comprises purchase information and/or sound

recording information corresponding to the sound recording most recently selected by

audio subsystem 102.

20   Video image generator 702 functions to create a video image based on the video

image specification contained in data packet 732. In one embodiment, after creating the

video image, generator 702 transmits the video image to transmission subsystem 190.

Transmission subsystem 190 functions to transmits the video image, data packet 731 (if

any), and the audio stream generated by audio subsystem 102 to transmission system 170. In one embodiment, the video image, data packet 731 and audio stream are transmitted together in an MPEG-2 data stream.

In the embodiment shown in FIG. 7, audio subsystem 102 performs process 300, as described above. However, video subsystem 104 does not perform process 330 and transmission system 170 does not perform process 360. Rather, video subsystem 104 performs process 800, which is shown in FIG. 8. Additionally, video image generator performs a process 900, which is shown in FIG. 9.

Process 800 begins in step 802, where video subsystem 104 waits for a trigger message from audio subsystem 102 or for a timer to expire. If video subsystem 104 receives a trigger message from audio subsystem 102, control passes to step 804, and if a timer expires, control passes to step 820.

In step 804, video subsystem 104 parses the trigger message to determine the sound recording identifier, sound recording information, and channel identifier specified therein. Next (step 806), video subsystem 104 uses this information, together with the pre-defined configuration data that is associated with the channel identified by the channel identifier, to create a data packet 731 for the identified channel.

Data packet 731 created in step 806 comprises purchase information and/or sound recording information. The purchase and/or sound recording information may be included in the trigger message and/or included in the pre-defined configuration data. After step 806, control passes to step 808. In step 808, video subsystem 104 uses the sound recording identifier determined in step 804 and the pre-defined configuration data to create a data packet 732. Data packet 732 comprises a video image specification

(e.g., a list of visual media asset identifiers together with their associated positions).

After generating data packets 731 and 732, video subsystem 104 performs steps 810 and

812. In step 810, video subsystem 104 transmits data packet 731 to transmission system

190 (or to transmission system 170). In step 812, video subsystem 104 provides data

5   packet 732 to video image generator 702.

In step 820, video subsystem 104 determines a channel and an asset identifier

queue that is associated with the expired timer. Next (step 822), video subsystem 104

creates for the identified channel a data packet 732 that comprises a video image

specification. Next (step 812) data packet 732 is provided to video image generator

10   702. After step 812, control passes back to step 802.

Referring now to process 900, process 900 begins in step 902, where video

image generator 702 waits to receive from video subsystem 104 a data packet 732,

which comprises a vide image specification. When a data packet 732 is received,

control passes to step 904, where video image generator 702 parses the video image

15   specification contained in the data packet 732 to determine the set of asset identifiers

specified therein and the screen positions associated with each asset identifier. After

step 904, control passes to step 906.

In step 906, video image generator 702 retrieves from storage 186 the visual

media assets identified by the asset identifiers determined in step 904. Alternatively, in

20   one embodiment, video image generator 702 does not have access to storage 186, but

video subsystem 104 does. In this embodiment, generator 702 requests video

subsystem 104 to retrieve and transmit to generator 702 the visual media assets

identified by the asset identifiers determined in step 904.

Next (step 908), video image generator 702 uses the retrieved visual media assets and the screen positions determined in step 904 to create a video image that conforms to the video image specification. Video image generator 702 then transmits the video image to transmission subsystem 190 (step 910). After step 910, control passes back to step 902.

In one embodiment, data packet 732 is an HTML document and video image generator 702 is a hardware/software device that convert the HTML document to an MPEG video presentation. In one specific embodiment, video image generator converts the HTML document into an MPEG I-frame followed by null P-frames. Such a device can be purchased from Liberate Technologies of San Carlos, CA.

FIG. 10 is a block diagram of a system 1000 for providing audio/video programming according to another embodiment of the present invention. System 1000 is similar to systems 100 and 700. However, in system 1000 video subsystem 104 comprises the video image generator 702, which may be implemented in hardware and/or software. In this embodiment, a data packet that comprises a video image specification, such as data packet 732, is not needed because video subsystem 104 itself creates the video images that compliment the audio service. FIG. 11 illustrates a process 1100 performed by video subsystem 104 according to the embodiment shown in FIG. 10.

Process 1100 begins in step 1102, where video subsystem 104 determines whether a trigger message from audio subsystem has been received. If a trigger message is received, control passes to step 1104, otherwise control passes to step 1103. In step 1103, video subsystem 104 determines whether an asset queue timer has

expired. If an asset queue timer expires, control passes to step 1140, otherwise control

passes back to step 1102.

In step 1104, video subsystem 104 parses the trigger message to determine the

sound recording identifier specified therein. Next (step 1106), video subsystem 104,

uses the pre-defined configuration data to determine a set of assets identifiers that are

associated with the sound recording identifier determined in step 1104. Video

subsystem 104 then determines the screen position that is associated with each asset

identifier in the set (step 1108). Next (step 1112), video subsystem 104 uses the pre-

defined configuration data to determine whether there are any asset identifier queues

associated with the sound recording identifier determined in step 1104. If there are,

control passes to step 1114, otherwise control passes to step 1128.

In step 1114, video subsystem 104 selects one of the queues that the

configuration data indicates is associated with the sound recording identifier. Next

(step 1116), video subsystem determines the asset identifier set in the selected queue

that is at the "head" of the selected queue. After step 1116, control passes to step 1118.

In step 1118, video subsystem 104 determines each asset identifier listed in the

asset identifier set determined in step 1116 together with each asset identifier's

associated screen position. Next (step 1120), video subsystem 104 determines the

duration associated with the asset identifier set determined in step 1116. Next (step

1122), video subsystem 104 activates the timer associated with the selected queue so

that the timer will expire after X amount of time has expired, where X is equal to the

duration determined in step 1118. After step 1122, control passes to step 1124.

In step 1124, video subsystem 104 determines whether there are additional asset

identifier queues associated with the sound recording identifier. If there are, control

passes to step 1126, otherwise control passes to step 1128. In step 1126, video

subsystem 104 selects a queue that is associated with the sound recording and that has

not already been selected. After step 1126, control passes back to step 1116.

In step 1128, video subsystem 104 retrieves the assets identified by the asset

identifiers determined in steps 1106 and 1118. Next (step 1130), video subsystem 104

creates a video image using the retrieved assets, wherein each asset is positioned in the

video image according its associated position. After step 1130, control passes to step

1130, where the video image is transmitted to transmission system 190. After step

1132, control passes back to step 1102.

In step 1140, video subsystem 104 determines the queue that is associated with

the timer that expired. Next (step 1142), video subsystem 104 increments the head

pointer associated with the queue determined in step 1140 to point to the next asset

identifier set in the queue if the queue determined in step 1140 is associated with the

sound recording identifier determined in step 1104. After step 1142, control passes to

step 1106.

In another embodiment, the video images that complement the audio service are

pre-generated. That is, they are generated prior to the time when they are scheduled to

be displayed. For example, they may be generated one day or one week prior to when

they are scheduled to be displayed.

In this embodiment where video images are pre-generated, a data structure (e.g.,

a configuration file) associates the sound recording identifiers listed in a playlist with an

ordered set of video image identifiers, where each video image identifier identifies a

pre-generated video image. The set may contain one or more video image identifiers. If the ordered set of video image identifiers associated with a sound recording identifier contains more than one video image identifier, then each video image identifier in the set, except the video image identifier that is last in the order, is associated with a time duration. The data structure may also associate purchase information with each sound recording identifier.

FIG. 12 illustrates an exemplary data structure 1200 that associates sound recording identifiers from a playlist with a set of one or more video image identifiers. For example, sound recording identifier 1202 is associated with an ordered set 1204 of video image identifiers and is associated with purchase information 1205.

The ordered set of video image identifiers 1204 includes video image identifiers 1210, 1211, and 1212. Additionally, each video image identifier in set 1204, except for video image identifiers 1212, which is the last video image identifier in the order, is associated with a time duration.

Either video subsystem 104 or transmission system 170 may be able to retrieve the pre-generated video images from the storage unit in which they are stored. Thus, for example, the pre-generated video images may be stored in storage unit 185 or storage unit 186. Similarly, either video subsystem 104 or transmission system 170 may be able to retrieve data structure 1200.

If, for example, the pre-generated video images are stored in storage unit 185 and transmission system 170 has access to data structure 1200, then the trigger message generated by audio subsystem 102 may be sent to transmission system 170 instead of to video subsystem 104. In this embodiment, transmission system 170 performs process

1300 (see FIG. 13).

Process 1300 begins in step 1302, where transmission system 170 receives a trigger message that includes a sound recording identifier. Next (step 1304) transmission system 170 parses the trigger message to determine the sound recording identifier included therein. Next (step 1305), transmission system 170 accesses data structure 1200 to determine the ordered set of video image identifiers and purchase information that are associated with the sound recording identifier determined in step 1304. Next (step 1306), transmission system 170 retrieves from storage unit 185 the video image identified by the first identifier in the set.

Next (step 1308), transmission system 170 determines, based on the purchase information (or lack thereof), whether it should overlay Buy button 250 on the video image or send a command to the receivers 180 that causes the receivers to overlay Buy button 205 on the video image. If it should, control passes to step 1310, otherwise control passes to step 1311. In step 1310, transmission system 170 transmits to receivers 180 the most recently retrieved video image with Buy button 250 included in the video image (or transmits to receivers 180 the video image together with a command that instructs receivers 180 to display Buy button 250). In step 1311, transmission system 170 transmits to receivers 180 the video image only.

Next (step 1312), transmission system 170 accesses data structure 1200 to determine whether there is a time duration associated with the video image transmitted in step 1310 or 1311. That is, transmission system 170 determines whether data structure 1200 associates a time duration with the video image identifier that identifies the video image. If there is no time duration associated with the video image, then

control passes back to step 1302, otherwise control passes to step 1314. In step 1314, transmission system 170 sets a timer to expire after X seconds and activates the timer, where X is the time duration in seconds associated with the video image transmitted in step 1310 or 1311. When the timer expires, transmission system 170 retrieves from storage unit 185 the video image identified by the next identifier in the set (step 1316) After step 1316, control passes back to step 1308.

If, for example, the pre-generated video images are stored in storage unit 185 but transmission system 170 does not have access to data structure 1200, then the trigger message is sent to video subsystem 104, which will have access to data structure 1200. In this embodiment, video subsystem 104 and transmission system 170 perform processes 1400 (see FIG. 14A) and process 1450 (see FIG. 14B), respectively. Alternatively, video subsystem 104 and transmission system 170 perform processes 1500 (see FIG. 15A and 1550 (see FIG. 15B), respectively.

Process 1400 begins in step 1402, where video subsystem 104 receives a trigger message that includes a sound recording identifier. Next (step 1404) video subsystem 104 parses the trigger message to determine the sound recording identifier included therein. Next (step 1406), video subsystem 104 accesses data structure 1200 to determine the ordered set of video image identifiers that is associated with the sound recording identifier determined in step 1404. Next (step 1407), video subsystem 104 selects the first video image identifier from the ordered set of video image identifiers.

Next (step 1408), video subsystem 104 transmits the most recently selected video image identifier to transmission system 170. In addition to transmitting the video image to transmission system 170, video subsystem may also transmit to transmission

system 170 purchase information and/or commands that instruct transmission system 170 to overlay selectable buttons (e.g., Buy button 250) on the video image to create an interactive service for the listeners. After step 1408, control passes to step 1410.

In step 1410, video subsystem 104 accesses data structure 1200 to determine whether there is a time duration associated with the video image identifier transmitted in step 1408. If there is no time duration associated with the video image identifier, then control passes back to step 1402, otherwise control passes to step 1414.

In step 1414, video subsystem 104 sets a timer to expire after X seconds and activates the timer, where X is the time duration in seconds associated with the video image identifier. When the timer expires, video subsystem 104 selects the next identifier in the ordered set (step 1416). After step 1416, control passes back to step 1408.

Process 1450 begins in step 1452, where transmission system 170 receives a video image identifier and purchase information (if any) from video subsystem 104. Next (step 1456), transmission system 170 retrieves from storage unit 185 the video image identified by the received identifier. Next (step 1458), transmission system 170 determines, based on the purchase information (or lack thereof), whether it should overlay Buy button 250 on the video image or send a command to the receivers 180 that causes the receivers to overlay Buy button 205 on the video image. If it should, control passes to step 1460, otherwise control passes to step 1461. In step 1460, transmission system 170 transmits to receivers 180 the retrieved video image with Buy button 250 included in the video image (or transmits to receivers 180 the video image together with a command that instructs receivers 180 to display Buy button 250). In step 1461,

transmission system 170 transmits to receivers 180 the video image only. After steps 1460 and 1461 control passes back to step 1452.

Process 1500 begins in step 1502, where video subsystem 104 receives a trigger message that includes a sound recording identifier. Next (step 1504) video subsystem 104 parses the trigger message to determine the sound recording identifier included therein. Next (step 1506), video subsystem 104 accesses data structure 1200 to determine the ordered set of video image identifiers that is associated with the sound recording identifier determined in step 1504. Next (step 1508), video subsystem 104 transmits to transmission system 170 the ordered set of video image identifiers and the purchase information associated with the sound recording identifier. After step 1508, control passes back to step 1502.

Process 1550 is similar to process 1300. Process 1550 begins in step 1552, where transmission system 170 receives the ordered set of video image identifiers and purchase information. After step 1552, transmission system 170 performs steps 1306-1316. After step 1316, control passes back to step 1552.

If, for example, the pre-generated video images are stored in storage unit 186 instead of 185 and video subsystem 104 has access to data structure 1200, then the trigger message generated by audio subsystem 102 is sent to video subsystem 104. In this embodiment, video subsystem 104 performs process 1600 (see FIG. 16).

Process 1600 begins in step 1602 where video subsystem 104 receives a trigger message that includes a sound recording identifier. Next (step 1604) video subsystem 104 parses the trigger message to determine the sound recording identifier included therein. Next (step 1606), video subsystem 104 accesses data structure 1200 to

determine the ordered set of video image identifiers that is associated with the sound recording identifier determined in step 1604. Next (step 1608), video subsystem 104 retrieves from storage unit 186 the video image identified by the first identifier in the set. Next (step 1610), video subsystem 104 transmits the most recently retrieved video image to transmission system 170. In addition to transmitting the video image to transmission system 170, video subsystem may also transmit to transmission system 170 purchase information and/or commands that instruct transmission system 170 to overlay selectable buttons (e.g., Buy button 250) on the video image to create an interactive service for the listeners. After step 1610, control passes to step 1612.

In step 1612, video subsystem 104 accesses data structure 1200 to determine whether there is a time duration associated with the video image transmitted in step 1610. That is, video subsystem 104 determines whether data structure 1200 associates a time duration with the video image identifier that identifies the video image. If there is no time duration associated with the video image, then control passes back to step 1602, otherwise control passes to step 1614. In step 1614, video subsystem 104 sets a timer to expire after X seconds and activates the timer, where X is the time duration in seconds associated with the video image. When the timer expires, video subsystem 104 retrieves from storage unit 186 the video image identified by the next identifier in the set (step 1616). After step 1616, control passes back to step 1610.

While various embodiments/variations of the present invention have been described above, it should be understood that they have been presented by way of example only, and not limitation. Thus, the breadth and scope of the present invention should not be limited by any of the above-described exemplary embodiments, but

should be defined only in accordance with the following claims and their equivalents.